

BAB II

LANDASAN TEORI

2.1 Statistika Deskriptif

Statistika deskriptif merupakan fase statistik yang melibatkan pengumpulan, penyusunan, pengolahan, dan penyajian data dalam bentuk tabel, grafik, dan diagram maupun peta untuk memberikan gambaran yang jelas dan ringkas tentang suatu keadaan atau peristiwa. Pada statistika deskriptif, peta digunakan untuk menampilkan data geografis atau spasial, memberikan gambaran visual tentang distribusi atau pola data di berbagai lokasi. Tekanan umumnya diberikan pada pengumpulan dan penataan data serta penggunaan pengukuran-pengukuran yang sifatnya merupakan penyederhanaan, misalnya rata-rata dan sebagainya yang dapat secara efektif menggambarkan karakteristik subjek yang diteliti (Dajan, 2008). Ukuran dispersi yang sering digunakan adalah rata-rata, varian, dan standar deviasi.

Tabel 2.1 Rumus Statistika Deskriptif

	Populasi	Sampel
Rata-rata	$\mu = \frac{\sum X}{N}$	$\bar{X} = \frac{\sum X}{n}$
Varian	$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$	$s^2 = \frac{\sum (X - \bar{X})^2}{n - 1}$
Standar Deviasi	$\sigma = \sqrt{\text{varian populasi}}$	$s = \sqrt{\text{varian sampel}}$

2.2 Standarisasi Data

Tujuan standarisasi data adalah untuk mengubah nilai-nilai variabel awal menjadi nilai-nilai baru dengan rata-rata nol dan simpangan baku satu. Ini dilakukan jika satuan data sangat berbeda untuk setiap variabel yang akan dipelajari. Dalam kasus ini, standarisasi data dilakukan dengan mengubah data saat ini ke nilai *Z-score* (Lind dkk., 2007). Persamaan untuk standarisasi adalah sebagai berikut.

$$Z = \frac{x_i - \bar{x}}{s} \quad (2.1)$$

Dimana :

Z : Nilai *Z-score*

x_i : Objek, $i = 1, 2, \dots, n$

$\bar{x}$: Nilai rata-rata variabel

s : Standar deviasi

2.3 Analisis Regresi Linear Berganda

Analisis regresi merupakan salah satu metode statistika yang digunakan dalam berbagai bidang. Analisis ini bertujuan untuk melihat hubungan dan memodelkan satu variabel dependen (Y) dengan satu atau lebih variabel independen (X). Secara umum hubungan tersebut dapat dinyatakan dalam bentuk model sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i$$

Dimana :

Y_i : Variabel dependen untuk pengamatan ke- i , untuk $i = 1, 2, \dots, n$

β_0 : Konstanta

$\beta_1, \beta_2, \dots, \beta_p$: Nilai koefisien variabel independen

$X_{1i}, X_{2i}, \dots, X_{pi}$: Variabel independen pada amatan ke- i

ε_i : Galat (*error*) untuk pengamatan ke- i

Dalam notasi matriks persamaan diatas dapat ditulis menjadi :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.2)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1p} \\ 1 & X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$\mathbf{y}$: Matriks variabel dependen berukuran $(n \times 1)$

$\mathbf{X}$: Matriks variabel independen berukuran $(n \times (p + 1))$

$\boldsymbol{\beta}$: Matriks parameter berukuran $((p + 1) \times 1)$

$\boldsymbol{\varepsilon}$: Matriks dari galat (*error*) berukuran $n \times 1$

2.4 Estimasi Parameter Regresi Linear Berganda

Tujuan dari estimasi parameter regresi linear berganda adalah untuk mendapatkan model regresi linear berganda yang akan digunakan untuk menganalisis model. Model yang digunakan pada regresi linear berganda adalah metode *Ordinary Least Square* (OLS). Metode ini bertujuan untuk meminimumkan jumlah kuadrat *error* (Widarjono, 2015). Dari persamaan (2.2) dapat ditulis matriks *error* sebagai berikut.

$$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \quad (2.3)$$

Sehingga diperoleh

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\ \sum_{i=1}^n \varepsilon_i^2 &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ \sum_{i=1}^n \varepsilon_i^2 &= \mathbf{y}^T \mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.4)$$

Karena $\mathbf{y}^T \mathbf{X}\boldsymbol{\beta}$ merupakan matriks yang berukuran 1×1 , maka matriksnya sama dengan matriks transposenya. Sehingga diperoleh persamaan (2.5) dibawah ini.

$$\mathbf{y}^T \mathbf{X}\boldsymbol{\beta} = (\mathbf{y}^T \mathbf{X}\boldsymbol{\beta})^T = \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} \quad (2.5)$$

Dengan mensubstitusikan persamaan (2.5) ke persamaan (2.4) maka diperoleh persamaan menjadi:

$$\begin{aligned} S = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} &= \mathbf{y}^T \mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}^T \mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}^T \mathbf{y} - 2\boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.6)$$

Selanjutnya untuk mendapatkan estimasi $\hat{\boldsymbol{\beta}}$, pada persamaan (2.6) akan di diferensialkan terhadap $\boldsymbol{\beta}^T$, sehingga diperoleh

$$\frac{\partial(\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})}{\partial \boldsymbol{\beta}^T} = 0$$

$$\frac{\partial(\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})}{\partial \boldsymbol{\beta}^T} = \frac{\partial(\mathbf{y}^T \mathbf{y} - 2\boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} = 0$$

$$\frac{\partial(\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})}{\partial \boldsymbol{\beta}^T} = 0 - 2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \boldsymbol{\beta} = 0$$

$$-2\mathbf{X}^T \mathbf{X} \boldsymbol{\beta} = -2\mathbf{X}^T \mathbf{y}$$

$$\mathbf{X}^T \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^T \mathbf{y}$$

$$\boldsymbol{\beta} = \frac{\mathbf{X}^T \mathbf{y}}{\mathbf{X}^T \mathbf{X}}$$

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Sehingga diperoleh estimator $\hat{\boldsymbol{\beta}}$ adalah

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (2.7)$$

2.5 Pengujian Parameter Regresi Linear Berganda

Pengujian parameter dalam model regresi linear berganda ini bertujuan untuk mengetahui apakah parameter tersebut telah menunjukkan hubungan yang nyata antara variabel dependen dan variabel independen serta untuk mengetahui kelayakan parameter dalam menerangkan model. Dalam pengujian parameter regresi, terdapat dua uji yang perlu dilakukan yaitu uji simultan dan uji parsial.

2.5.1 Uji Simultan Regresi Linear Berganda

Pengujian parameter secara simultan adalah uji semua variabel independen secara keseluruhan bersamaan di dalam suatu model. Uji simultan atau yang biasa disebut uji F bertujuan untuk menentukan apakah variabel independen secara bersama-sama berpengaruh signifikan terhadap variabel dependen.

Hipotesis yang digunakan adalah :

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_p = 0$$

$$H_1 : \text{Paling sedikit ada satu } \beta_k \neq 0, k = 1, 2, \dots, p$$

Hipotesis tersebut menunjukkan bahwa untuk hipotesis nol seluruh variabel independen secara bersamaan tidak berpengaruh signifikan terhadap variabel dependen. Untuk hipotesis alternatifnya menunjukkan bahwa variabel independen

bersama-sama berpengaruh signifikan terhadap variabel dependen atau minimal ada satu variabel independen yang berpengaruh signifikan terhadap variabel dependen. Pengambilan keputusan dapat diketahui dengan menggunakan rumus uji F sebagai berikut :

$$F_{hit} = \frac{KTR}{KTE}$$

Pengujian diatas dapat diperoleh dengan menggunakan analisis varian atau dalam bahasa inggris disebut *Analysis of Variance* (ANOVA) (Kurniawan & Yuniarto, 2016), seperti dalam Tabel 2.2 berikut.

Tabel 2.2 Tabel ANOVA

Sumber Keberagaam	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F _{hitung}
Regresi	p	$JKR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$KTR = \frac{JKR}{p}$	$\frac{KTR}{KTE}$
Error (Galat)	$n - p - 1$	$JKE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$	$KTE = \frac{JKE}{n - p - 1}$	
Total	$n - 1$	$JKT = \sum_{i=1}^n (y_i - \bar{y})^2$		

Dimana:

$\hat{y}_i$: Nilai estimasi variabel dependen pengamatan ke- i

$\bar{y}$: Nilai rata-rata variabel dependen

p : Banyaknya variabel independen

n : Banyaknya pengamatan yang diteliti

α : Tingkat kepercayaan (*alpha*)

JKR : Jumlah Kuadrat Regresi

JKE : Jumlah Kuadrat Error

JKT : Jumlah Kuadrat Total

KTR : Kuadrat Tengah Regresi

KTE : Kuadrat Tengah *Error*

Kriteria pengujian untuk uji simultan yaitu :

1. Jika $F_{hit} > F_{tabel}(F_{(a;p;n-p-1)})$ dan nilai $p-value \leq a$ maka tolak H_0
2. Jika $F_{hit} \leq F_{tabel}(F_{(a;p;n-p-1)})$ dan nilai $p-value > a$ maka terima H_0

Apabila keputusannya adalah tolak H_0 , dapat disimpulkan bahwa sekurang-kurangnya ada satu variabel yang berpengaruh secara signifikan.

2.5.2 Uji Parsial Regresi Linear Berganda

Uji parsial merupakan pengujian secara individu yang digunakan untuk menguji ada atau tidaknya pengaruh masing-masing variabel independen terhadap variabel dependen. Dalam regresi berganda, hal ini perlu dilakukan karena setiap variabel independen memberi pengaruh yang berbeda dalam model.

Hipotesis yang digunakan adalah :

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, \text{ dengan } k = 1, 2, \dots, p$$

Statistik uji yang digunakan dalam pengujian parsial adalah:

$$t_{hit} = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)} = \frac{\hat{\beta}_k}{\sqrt{var(\hat{\beta}_k)}}$$

Dimana $\hat{\beta}_k$ adalah nilai estimasi β_k (yang diperoleh dari metode OLS) dan $se(\hat{\beta}_k)$ adalah *standard error* bagi β_k .

Keputusan yang diambil dapat diperoleh dari perbandingan nilai t_{hit} dengan nilai $t_{tabel} \left(t_{\left(\frac{a}{2}, n-p-1 \right)} \right)$ atau bisa juga dilakukan dengan melihat nilai signifikansinya. Kriteria pengambilan keputusannya adalah tolak H_0 ketika $t_{hit} > t_{\left(\frac{a}{2}, n-p-1 \right)}$ atau $p-value \leq \alpha$, artinya variabel independen (X_i) memberikan pengaruh yang signifikan terhadap variabel dependen (Y). Sehingga dari uji ini dapat diketahui pengaruh tiap-tiap variabel independen terhadap variabel dependennya (Kurniawan & Yuniarto, 2016).

2.6 Uji Asumsi Klasik

Uji asumsi klasik merupakan uji data yang digunakan untuk mengetahui apakah data penelitian memenuhi syarat untuk dianalisis lebih lanjut, guna memberi kepastian bahwa persamaan regresi yang didapatkan memiliki ketepatan dalam estimasi, tidak bias, dan konsisten. Jenis pengujian asumsi klasik juga disesuaikan dengan teknik analisis yang digunakan. Uji asumsi klasik yang digunakan dalam penelitian ini terdiri dari uji multikolinearitas, uji normalitas, dan uji heteroskedastisitas.

2.6.1 Uji Multikolinearitas

Multikolinearitas merupakan hubungan linear antar variabel independen di dalam regresi linear berganda. Fungsi dari uji ini adalah untuk melihat ada atau tidaknya suatu korelasi yang tinggi antar variabel independen. Salah satu syarat yang harus dipenuhi dalam model regresi adalah tidak adanya hubungan antar variabel independen atau disebut tidak terjadi multikolinearitas. Jika terdapat korelasi yang tinggi di antara variabel-variabel independen, dapat menyebabkan masalah multikolinearitas.

Salah satu metode yang dapat digunakan untuk mendeteksi ada atau tidaknya multikolinearitas dalam suatu model regresi adalah dengan *Variance Inflation Factor* (VIF) (Kurniawan & Yuniarto, 2016). Pada umumnya jika VIF lebih besar dari 10, maka dapat dikatakan bahwa terdapat multikolinearitas antar variabel independen dengan rumus sebagai berikut.

$$VIF = \frac{1}{(1 - R_j^2)} \quad (2.8)$$

Dimana R_j^2 adalah koefisien determinasi dari regresi antar variabel independen ke- j .

$$R_j^2 = 1 - \frac{\sum_{i=1}^n (X_i - \hat{X}_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (2.9)$$

Kriteria pengujian :

1. Jika nilai $VIF < 10$ maka tidak terdapat multikolinearitas dalam model regresi linear berganda.
2. Jika nilai $VIF \geq 10$ maka terdapat multikolinearitas dalam model regresi linear berganda.

Jika dalam pengujian asumsi terdapat multikolinearitas, ada beberapa alternatif cara untuk mengatasi multikolinearitas seperti mengganti atau mengeluarkan variabel yang mempunyai korelasi yang tinggi. Namun hal ini harus dilakukan secara hati-hati, agar tidak menyebabkan variabel yang dikeluarkan merupakan variabel yang penting. Dengan menambahkan jumlah data (observasi) diharapkan dapat mengatasi multikolinearitas. Sampel yang lebih besar selalu akan meningkatkan tingkat akurasi atau presisi estimasi parameter. Mentransformasikan data ke dalam bentuk lain, misalnya logaritma natural, akar kuadrat, atau bentuk *first difference* dan *ratio transformation method* juga merupakan alternatif cara untuk mengatasi multikolinearitas (Kurniawan & Yuniarto, 2016).

2.6.2 Uji Normalitas Residual

Menurut Ghozali (2016), uji Normalitas bertujuan untuk mengetahui apakah masing-masing variabel berdistribusi normal atau tidak. Model regresi yang baik adalah memiliki nilai residual yang berdistribusi Normal. Uji normalitas dapat dideteksi dengan melihat grafik normal *probability plot* (Q-Q plot), jika data menyebar disekitar garis diagonal dan mengikuti arah garis diagonal menunjukkan pola distribusi normal, maka model regresi memenuhi asumsi normalitas. Bila data menyebar jauh dari garis diagonalnya atau tidak mengikuti arah garis diagonal, maka model regresi tidak memenuhi asumsi. Selain dengan melihat grafik, salah satu metode lainnya yang dapat digunakan untuk pengujian residual berdistribusi Normal adalah dengan menggunakan uji *Kolmogorov-Smirnov* (K-S). Pengujian statistik untuk uji *Kolmogorov-Smirnov* adalah sebagai berikut :

Hipotesis:

H_0 : Residual (*error*) berdistribusi Normal

H_1 : Residual (*error*) tidak berdistribusi Normal

Statistik uji:

$$D = \max_x |F_n(x) - F_0(x)| \quad (2.10)$$

Dimana:

$F_n(x)$: Distribusi frekuensi kumulatif empiris atau yang sedang diuji

$F_0(x)$: Distribusi frekuensi kumulatif teoritis yang dinyatakan dalam hipotesis nol

Untuk menghitung $F_0(x_i)$ dibantu tabel distribusi normal baku Z , dengan $Z = \frac{x_i - \bar{x}}{s}$.

$\bar{x}$ ialah rata-rata dan s ialah simpangan baku.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$
$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|^2}$$

Dengan taraf signifikansi sebesar α , maka kriteria pengujian sebagai berikut :

1. Jika $p\text{-value} \geq 0,05$ maka terima H_0 , artinya *error* data berdistribusi Normal.
2. Jika $p\text{-value} < 0,05$ maka tolak H_0 , artinya *error* data tidak berdistribusi Normal.

Cara untuk mengatasi data yang tidak berdistribusi Normal yaitu dengan melakukan transformasi data. Transformasi dilakukan bukan berarti data yang digunakan diubah dari bentuk aslinya, tetapi hanya diubah agar memenuhi asumsi-asumsi analisis ragam. Bentuk transformasi data untuk mengatasi normalitas, yaitu seperti transformasi log, transformasi akar, dan transformasi Arcsin.

Transformasi log digunakan untuk data asli kurang dari 10 atau banyaknya mendekati nol, maka dapat ditransformasikan dengan $\log(x + 1)$ atau dapat menggunakan $\log(10x)$ karena cenderung mendekati nol. Transformasi akar biasa digunakan ketika data tidak memenuhi kehomogenan. Transformasi dilakukan dengan rumus $\sqrt{x + 0.5}$ untuk data yang < 10 . Jika data berbentuk proporsi atau persentase, maka transformasi Arcsin dapat digunakan (Kurniawan & Yuniarto, 2016).

2.6.3 Uji Heteroskedastisitas

Menurut Ghozali (2016), uji heteroskedastisitas bertujuan untuk melihat apakah terdapat kesamaan suatu varian dari residual (*error*) pengamatan satu ke pengamatan yang lainnya. Jika suatu varians dari residual satu pengamatan ke pengamatan yang lain bersifat tetap, maka dapat disebut homoskedastisitas (sama). Salah satu asumsi yang harus terpenuhi dalam model regresi linear berganda adalah varians variabel residual bersifat homoskedastisitas. Untuk menguji ada atau tidaknya homoskedastisitas dalam regresi dapat dilakukan dengan menggunakan beberapa cara, salah satunya adalah menggunakan uji *Glejser*. Uji ini digunakan untuk mengetahui apakah sebuah model regresi memiliki indikasi heteroskedastisitas dengan cara meregresi nilai absolut *error*. Heteroskedastisitas terjadi apabila hasil regresi nilai absolut *error* terhadap variabel memiliki nilai signifikansi $< 0,05$. Pengujian statistik untuk uji *Glesjer* adalah sebagai berikut :

Hipotesis :

$H_0 : var(\varepsilon_i) = \sigma^2$ (variens *error* bersifat homoskedastisitas)

$H_1 : var(\varepsilon_i) \neq \sigma^2$ (variens *error* bersifat heteroskedastisitas)

Statistik uji :

$$|\hat{u}_i| = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_p X_p \quad (2.11)$$

Dimana $|\hat{u}_i|$ merupakan nilai absolut dari nilai *error* pada regresi berganda metode OLS. Dengan kriteria pengujian sebagai berikut :

1. Jika nilai *p-value* $\geq 0,05$ maka H_0 diterima yang berarti variansi *error* bersifat homoskedastisitas.
2. Jika nilai *p-value* $< 0,05$ maka H_0 ditolak yang berarti variansi *error* bersifat heteroskedastisitas.

Pada pengujian asumsi heteroskedastisitas, jika terindikasi adanya heteroskedastisitas dapat diatasi dengan melakukan transformasi logaritma normal. Galat (*error*) suatu data yang ditransformasi logaritma normal akan mengecil. Karena *error*-nya mengecil, akibatnya heteroskedastisitas akan berkurang (Kurniawan & Yuniarto, 2016).

2.7 Analisis Regresi Spasial

Regresi spasial merupakan pengembangan teori dari regresi linear klasik. Regresi spasial pertama kali diperkenalkan oleh Anselin (1988) berdasarkan hukum Tobler I yang menyatakan bahwa segala sesuatu saling berhubungan satu dengan yang lain, tetapi sesuatu yang dekat memiliki pengaruh yang lebih besar daripada sesuatu yang jauh.

Analisis regresi spasial adalah analisis yang mengevaluasi hubungan variabel dependen dengan beberapa variabel independen dengan memberikan efek spasial pada beberapa lokasi yang menjadi pusat pengamatan. Model umum regresi spasial atau juga yang biasa disebut *Spasial Autoregressive Moving Average* (SARMA) dinyatakan sebagai berikut :

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (2.12)$$

Dengan

$$\begin{aligned} \mathbf{u} &= \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &\sim N(0, \sigma^2 \mathbf{I}) \end{aligned} \quad (2.13)$$

Dimana:

- $\mathbf{y}$: Vektor variabel dependen berukuran $n \times 1$
- ρ : Parameter koefisien spasial lag variabel dependen
- $\mathbf{W}\mathbf{y}$: Matriks pembobot spasial berukuran $n \times n$
- $\mathbf{X}$: Matriks variabel independen berukuran $n \times (p + 1)$
- $\boldsymbol{\beta}$: Vektor parameter koefisien regresi berukuran $(p + 1) \times 1$
- λ : Parameter koefisien spasial lag pada *error*
- $\mathbf{u}$: Vektor *error* pada persamaan (2.12) berukuran $n \times 1$
- $\boldsymbol{\varepsilon}$: Vektor *error* pada persamaan (2.13) yang diasumsikan berdistribusi normal dengan mean nol dan varians $\sigma^2 \mathbf{I}$
- $\mathbf{I}$: Matriks identitas dengan ukuran $n \times n$

Pengujian asumsi pada regresi spasial sama halnya dengan pengujian asumsi pada regresi linear berganda. Pengujian asumsi tersebut adalah asumsi kehomogenan, kenormalan dan asumsi tidak ada multikolinearitas antar variabel independen.

2.8 Spatial Autoregressive (SAR)

Analisis regresi spasial lag mengasumsikan bahwa nilai variabel di suatu lokasi dipengaruhi oleh nilai variabel di lokasi tetangganya. Dengan memperhitungkan hubungan spasial antara lokasi-lokasi yang berdekatan, digunakan matriks ketergantungan spasial (matriks *contiguity*). *Spatial Autoregressive* (SAR) atau disebut juga *Spatial Lag Model* (SLM) merupakan salah satu model spasial pendekatan area dengan mengkombinasikan model regresi sederhana dengan spasial lag pada variabel dependen dengan menggunakan data *cross section* (Anselin, 1988). Jika $\rho \neq 0$ dan $\lambda = 0$ maka model tersebut merupakan model *Spatial Autoregressive*. Bentuk umum SAR adalah sebagai berikut:

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &\sim N(0, \sigma^2 \mathbf{I}) \end{aligned} \quad (2.14)$$

Dimana:

- $\mathbf{y}$: Vektor variabel dependen berukuran $n \times 1$
- ρ : Parameter koefisien spasial lag variabel dependen, parameter yang mengukur tingkat korelasi pengaruh spasial dari suatu wilayah terhadap wilayah lain disekitarnya.
- $\mathbf{W}\mathbf{y}$: Matriks pembobot spasial variabel dependen berukuran $n \times n$
- $\boldsymbol{\beta}$: Vektor parameter koefisien regresi berukuran $(p + 1) \times 1$
- $\mathbf{X}$: Matriks variabel independen berukuran $n \times (p + 1)$
- $\boldsymbol{\varepsilon}$: Vektor *error* berukuran $n \times 1$
- n : Jumlah lokasi pengamatan
- $\mathbf{I}$: Matriks identitas dengan ukuran $n \times n$

Persamaan (2.14) dapat diekspresikan menjadi :

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \mathbf{y} - \rho \mathbf{W}\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ (\mathbf{I} - \rho \mathbf{W})\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= (\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.15)$$

Turunan dari persamaan (2.15) terhadap $\mathbf{y}$ akan membentuk fungsi *Jacobian* (J) seperti berikut :

$$J = \left| \frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{y}} \right| = |\mathbf{I} - \rho \mathbf{W}|$$

Estimasi parameter model SAR dengan menggunakan MLE, membentuk fungsi *likelihood* (L) dari $\boldsymbol{\varepsilon}$ seperti pada persamaan (2.16) dan fungsi *likelihood* dari variabel y seperti pada persamaan (2.17).

$$L(\sigma^2; \boldsymbol{\varepsilon}) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left(-\frac{1}{2\sigma^2} (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}) \right) \quad (2.16)$$

$$L(\rho, \beta, \sigma^2 | y) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} (J) \exp \left(-\frac{1}{2\sigma^2} (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}) \right) \quad (2.17)$$

Dengan mensubstitusi persamaan (2.15) dan fungsi *Jacobian* pada persamaan (2.17) akan menghasilkan persamaan (2.18) sebagai berikut :

$$L(\rho, \beta, \sigma^2 | y) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} |\mathbf{I} - \rho \mathbf{W}| \exp \left(-\frac{1}{2\sigma^2} \left[((\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T ((\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \right) \quad (2.18)$$

Sehingga dengan memaksimalkan fungsi pada persamaan (2.18), operasi logaritma natural *likelihood* maka didapatkan persamaan berikut :

$$\begin{aligned} \ln(L) &= \frac{n}{2} \ln \left(\frac{1}{2\pi\sigma^2} \right) + \ln |\mathbf{I} - \rho \mathbf{W}| - \frac{1}{2\sigma^2} \left[((\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T ((\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \\ \ln(L) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) + \ln |\mathbf{I} - \rho \mathbf{W}| - \frac{1}{2\sigma^2} \left[((\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T ((\mathbf{I} - \rho \mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \end{aligned} \quad (2.19)$$

Pada model SAR ini pendugaan parameternya menggunakan metode kemungkinan maksimum (*Maximum Log Likelihood*) dimana setiap parameter yang belum diketahui diperoleh dengan memaksimumkan suatu fungsi kemungkinan (*likelihood function*) (Yasin dkk., 2020).

2.8.1 Estimasi Parameter *Spatial Autoregressive Model (SAR)*

Estimasi untuk parameter σ^2 dan β diperoleh dengan memaksimumkan fungsi *log likelihood* pada persamaan (2.19). Estimator untuk σ^2 diperoleh melalui turunan persamaan (2.19) terhadap σ^2 .

$$\begin{aligned}\frac{\partial \ln(L)}{\partial \sigma^2} &= 0 \\ \frac{\partial \ln(L)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{n}{2(\sigma^2)^2} \left[((I - \rho W)y - X\beta)^T ((I - \rho W)y - X\beta) \right] \\ 0 &= -\frac{n}{2\sigma^2} + \frac{n}{2(\sigma^2)^2} \left[((I - \rho W)y - X\beta)^T ((I - \rho W)y - X\beta) \right] \\ \sigma^2 &= \frac{((I - \rho W)y - X\beta)^T ((I - \rho W)y - X\beta)}{n}\end{aligned}$$

Sehingga didapatkan estimator parameter untuk σ^2 sebagai berikut :

$$\hat{\sigma}^2 = \frac{((I - \rho W)y - X\beta)^T ((I - \rho W)y - X\beta)}{n} \quad (2.20)$$

Estimasi untuk parameter β diperoleh dengan cara yang sama yaitu memaksimumkan *ln likelihood* melalui turunan persamaan (2.19) terhadap β .

$$\begin{aligned}\frac{\partial \ln(L)}{\partial \beta} &= 0 \\ \frac{\partial \ln(L)}{\partial \beta} &= \frac{\partial \left\{ -\frac{1}{2\sigma^2} \left[((I - \rho W)y - X\beta)^T ((I - \rho W)y - X\beta) \right] \right\}}{\partial \beta} \\ 0 &= \frac{\partial \left\{ -\frac{1}{2\sigma^2} \left[((I - \rho W)y - X\beta)^T ((I - \rho W)y - X\beta) \right] \right\}}{\partial \beta} \\ 0 &= [X^T (I - \rho W)y - X^T \beta] \\ \beta &= (X^T X)^{-1} X^T (I - \rho W)y\end{aligned}$$

Sehingga didapatkan estimator parameter untuk β sebagai berikut :

$$\hat{\beta} = (X^T X)^{-1} X^T (I - \rho W)y \quad (2.21)$$

2.9 Spatial Error Model (SEM)

Spatial Error Model (SEM) merupakan model regresi spasial dimana pada peubah *error*nya terdapat korelasi spasial. Hal ini disebabkan adanya variabel independen yang tidak dilibatkan dalam model sehingga akan dihitung sebagai *error* dan variabel tersebut memiliki hubungan dengan *error* pada lokasi lain. Model SEM terbentuk apabila $\rho = 0$ dan $\lambda \neq 0$, sehingga model ini mengasumsikan bahwa proses *autoregressive* hanya pada *error* model (LeSage & Pace, 2009). Bentuk umum persamaan SEM adalah :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (2.22)$$

$$\mathbf{u} = (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon} \quad (2.23)$$

$$\boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$$

Persamaan (2.23) diperoleh penjabaran

$$\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}$$

$$\mathbf{u} - \lambda\mathbf{W}\mathbf{u} = \boldsymbol{\varepsilon}$$

$$(\mathbf{I} - \lambda\mathbf{W})\mathbf{u} = \boldsymbol{\varepsilon}$$

Sehingga persamaan (2.22) dan (2.23) dapat dinyatakan dengan :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon} \quad (2.24)$$

Dimana :

$\mathbf{y}$: Vektor variabel dependen berukuran $n \times 1$

λ : Parameter yang mengukur tingkat korelasi pengaruh *error* spasial dari suatu wilayah terhadap wilayah lain disekitarnya.

$\mathbf{W}$: Matriks pembobot spasial *error* berukuran $n \times n$. Matriks ini menggambarkan hubungan spasial antar wilayah yang bertetanggan.

$\boldsymbol{\varepsilon}$: Vektor *error* dengan konstanta variansi σ^2

Menurut Anselin (1988), metode *Maximum Likelihood Estimation* (MLE) digunakan untuk menduga parameter SEM. Dari persamaan (2.24) dibentuk fungsi *likelihood*, pembentuk fungsi *likelihood* tersebut dilakukan melalui residual ($\boldsymbol{\varepsilon}$). Hasil pembentukan fungsi residual tersebut yaitu :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} = \mathbf{y}(\mathbf{I} - \lambda\mathbf{W}) - (\mathbf{I} - \lambda\mathbf{W})\mathbf{X}\boldsymbol{\beta} \quad (2.25)$$

$$\boldsymbol{\varepsilon} = (\mathbf{I} - \lambda \mathbf{W})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Turunan dari persamaan (2.25) terhadap $\mathbf{y}$ akan membentuk fungsi *Jacobian*, seperti berikut :

$$\mathbf{J} = \left| \frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{y}} \right| = |\mathbf{I} - \lambda \mathbf{W}|$$

Estimasi parameter model SEM dengan menggunakan MLE, membentuk fungsi *likelihood* (L) dari $\boldsymbol{\varepsilon}$ seperti pada persamaan (2.26) dan fungsi *likelihood* dari variabel $\mathbf{y}$ seperti pada persamaan (2.27).

$$L(\sigma^2; \boldsymbol{\varepsilon}) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \exp \left(-\frac{1}{2\sigma^2} (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}) \right) \quad (2.26)$$

$$L(\lambda, \beta \sigma^2 | \mathbf{y}) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} (\mathbf{J}) \exp \left(-\frac{1}{2\sigma^2} (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}) \right) \quad (2.27)$$

Dari substitusi persamaan (2.25) dan fungsi *Jacobian* ke persamaan (2.27) yang akan menghasilkan persamaan (2.28) sebagai berikut :

$$L(\beta, \lambda, \sigma^2 | \mathbf{y}) = \left(\frac{1}{(2\pi)^{\frac{n}{2}} \sigma^2} \right) \mathbf{y} e^{-\frac{1}{2}(\mathbf{I}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{I} - \lambda \mathbf{W})^T (\sigma^2 \mathbf{I})^{-1} (\mathbf{I} - \lambda \mathbf{W})(\mathbf{I}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})} \quad (2.28)$$

Operasi logaritma natural *likelihood* pada persamaan (2.28) sehingga menjadi persamaan (2.29).

$$\ln(L) = \ln \left(\frac{1}{(2\pi)^{\frac{n}{2}} \sigma^2} \right) \mathbf{y} e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{I} - \lambda \mathbf{W})^T (\sigma^2 \mathbf{I})^{-1} (\mathbf{I} - \lambda \mathbf{W})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}$$

$$\begin{aligned} \ln(L) = & -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma^2) + \ln |\mathbf{I} - \lambda \mathbf{W}| \\ & - \frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{I} - \lambda \mathbf{W})^T (\mathbf{I} - \lambda \mathbf{W})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned} \quad (2.29)$$

2.9.1 Estimasi Parameter *Spatial Error Model* (SEM)

Estimasi parameter β dan σ^2 dapat diperoleh dengan cara memaksimumkan fungsi $\ln$ *likelihood* pada persamaan (2.29). Didapatkan estimasi parameter β sebagai berikut :

$$\begin{aligned}\frac{\partial \ln(L)}{\partial (\beta)^T} &= 0 \\ \frac{\partial \ln(L)}{\partial (\beta)^T} &= \frac{\partial \left(\frac{1}{2\sigma^2} (y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta) \right)}{\partial (\beta)^T} \\ 0 &= -\frac{1}{2\sigma^2} \frac{\partial [(y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta)]}{\partial (\beta)^T} \\ 0 &= \frac{1}{2\sigma^2} [(X^T (I - \lambda W)^T (I - \lambda W) X)^{-1} X^T (I - \lambda W)^T (I - \lambda W) y - \beta] \\ \beta &= (X^T (I - \lambda W)^T (I - \lambda W) X)^{-1} X^T (I - \lambda W)^T (I - \lambda W) y\end{aligned}$$

Sehingga didapatkan estimator parameter untuk β sebagai berikut :

$$\hat{\beta} = [(X^T - \lambda W X^T)^T (X - \lambda W X)]^{-1} (X^T - \lambda W X^T)^T (I - \lambda W) y \quad (2.30)$$

Estimator untuk σ^2 diperoleh dengan menurunkan persamaan $\ln$ *likelihood* (2.29) terhadap σ^2 adalah sebagai berikut.

$$\begin{aligned}\frac{\partial \ln(L)}{\partial \sigma^2} &= 0 \\ \frac{\partial \ln(L)}{\partial \sigma^2} &= \frac{-\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) + \ln|I - \lambda W| - \frac{1}{2\sigma^2} (y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta)}{\partial \sigma^2} \\ 0 &= -\frac{n}{2\sigma^2} + \frac{n}{2(\sigma^2)^2} [(y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta)] \\ 0 &= \frac{-n\sigma^2 + (y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta)}{2(\sigma^2)^2} \\ \frac{2\sigma^4 \cdot n\sigma^2}{2\sigma^4} &= (y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta)\end{aligned}$$

Sehingga didapatkan estimator parameter untuk σ^2 sebagai berikut :

$$\hat{\sigma}^2 = \frac{(y - X\beta)^T (I - \lambda W)^T (I - \lambda W) (y - X\beta)}{n} \quad (2.31)$$

2.10 Uji Dependensi Spasial

Dependensi spasial menggambarkan adanya hubungan fungsional yang terjadi pada suatu titik dalam ruang dan apa yang terjadi di tempat lain. Uji untuk mengetahui dependensi spasial dalam suatu model adalah dengan menggunakan *Moran's I* dan *Lagrange Multiplier (LM)* (Anselin, 1988).

2.10.1 Indeks *Moran's I*

Salah satu statistik umum yang digunakan dalam penelitian ini adalah statistik *Moran's I*. Indeks *Moran's I* adalah ukuran dari korelasi antara pengamatan yang saling berdekatan. Statistik ini membandingkan nilai pengamatan di daerah lainnya. Indeks *Moran's I* mengukur korelasi satu variabel independen misal X (x_i dan x_j) dimana $i \neq j, i = 1, 2, \dots, n; j = 1, 2, \dots, n$ merupakan indeks untuk wilayah pengamatan. Indeks *Moran's I* dapat diukur dengan persamaan berikut (Samadi dkk., 2017) :

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})^2} \quad (2.32)$$

Dimana:

n : Banyaknya pengamatan

$\bar{x}$: Nilai rata-rata dari x_i dari n lokasi

x_i : Nilai amatan pada lokasi ke- i

x_j : Nilai amatan pada lokasi ke- j

W_{ij} : Elemen matriks pembobot spasial baris ke- i kolom ke- j

Nilai dari indeks I ini berkisaran antara -1 dan 1. Identifikasi pola menggunakan kriteria nilai indeks I , jika $I > I_0$, maka mempunyai pola mengelompok (terdapat autokorelasi positif), jika $I = I_0$, mempunyai pola menyebar tidak merata (tidak ada autokorelasi), dan jika $I < I_0$, memiliki pola menyebar (terdapat autokorelasi negatif). I_0 adalah nilai harapan dari I dengan rumus:

$$E(I) = I_0 = -\frac{1}{n-1} \quad (2.33)$$

Pengujian hipotesis terhadap parameter I dapat dilakukan sebagai berikut :

$H_0 : I = 0$ (tidak ada autokorelasi antar lokasi)

$H_1 : I \neq 0$ (terdapat autokorelasi antar lokasi)

Dengan nilai statistik uji sebagai berikut:

$$Z_{hit} = \frac{I - I_0}{\sqrt{var(I)}} \quad (2.34)$$

Nilai statistik uji dimana Z_{hit} mengikuti sebaran normal. Pengambilan keputusan tolak H_0 jika $|Z_{hit}| > \frac{Z_\alpha}{2}$.

Dimana nilai $var(I)$ dari *Moran's I* adalah:

$$var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - [E(I)]^2 \quad (2.35)$$

Dengan:

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n W_{ij} ; \quad S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (W_{ji} + W_{ij})^2$$

$$S_2 = \sum_{i=1}^n \left(\sum_{j=1}^n W_{ij} + \sum_{i=1}^n W_{ji} \right)^2$$

Secara visual, *Moran's I* dapat dilihat dengan menggunakan *scatter plot*. *Moran's I scatter plot* bertujuan untuk melihat hubungan antar nilai amatan dalam vektor variabel dependen (deviate dari rata-rata) dan rata-rata nilai pengamatan tetangga dalam bentuk vektor Wy (LeSage & Pace, 2009).

Kuadran II <i>Low-High</i>	Kuadran I <i>High-High</i>
Kuadran III <i>Low-Low</i>	Kuadran IV <i>High-Low</i>

Gambar 2.1 *Moran's I Scatter Plot*

Berdasarkan Gambar 2.1 terdapat empat kuadran kartesian dalam *Moran's I scatter plot* yang berpusat pada nilai nol untuk sumbu horizontal dan vertikal. Keempat kuadran tersebut adalah :

- a. Kuadran I (*High-high*), menunjukkan suatu daerah dengan nilai amatan tinggi yang berbatasan dengan daerah lain (tetangga) yang memiliki nilai amatan tinggi.
- b. Kuadran II (*Low-High*), menunjukkan suatu daerah dengan nilai amatan rendah yang berbatasan dengan daerah lain (tetangga) yang memiliki nilai amatan tinggi.
- c. Kuadran III (*Low-Low*), menunjukkan suatu daerah dengan nilai amatan rendah yang berbatasan dengan daerah lain (tetangga) yang memiliki nilai amatan rendah.
- d. Kuadran IV (*High-Low*), menunjukkan suatu daerah dengan nilai amatan tinggi yang berbatasan dengan daerah lain (tetangga) yang memiliki nilai amatan tinggi.

2.10.2 Uji *Lagrange Multiplier* (LM)

Uji *Lagrange Multiplier* (LM) digunakan untuk memilih model regresi spasial yang sesuai. Uji *Lagrange Multiplier* terdiri dari LM_{lag} dan LM_{error} . Apabila LM_{lag} signifikan maka model yang sesuai adalah SAR, jika LM_{error} signifikan maka model yang sesuai adalah SEM.

Uji LM digunakan untuk menentukan kehadiran efek spasial atau tidak didalam model. Bentuk uji LM adalah sebagai berikut (Rati dkk., 2013):

- SAR

Hipotesis yang digunakan Uji LM_{lag} adalah:

$H_0 : \rho = 0$ (tidak ada efek spasial pada lag)

$H_1 : \rho \neq 0$ (terdapat efek spasial pada lag)

Statistik uji:

$$LM_{lag} = \frac{\left(\frac{\varepsilon^T W y}{s^2} \right)^2}{\left(\frac{(W X \beta)^T M (W X \beta) + t s^2}{s^2} \right)} \quad (2.36)$$

Dimana:

$M : I - X(X^t X)^{-1} X^t$

$t : tr[(W^T + W)W]$

$$S^2 : \frac{\varepsilon^T \varepsilon}{n}$$

Dalam mengambil keputusan, tolak H_0 jika $LM_{lag} > \chi^2_{(\alpha,1)}$ atau $p\text{-value} < \alpha$.

Jika LM_{lag} signifikan maka model yang sesuai adalah SAR.

- SEM

Hipotesis yang digunakan LM_{error} adalah:

$H_0 : \lambda = 0$ (tidak ada efek spasial pada *error*)

$H_1 : \lambda \neq 0$ (terdapat efek spasial pada *error*)

Statistik uji:

$$LM_{error} = \frac{\left(\frac{\varepsilon^T W \varepsilon}{S^2} \right)^2}{t} \quad (2.37)$$

Dalam mengambil keputusan, tolak H_0 jika $LM_{error} > \chi^2_{(\alpha,1)}$ atau $p\text{-value} < \alpha$ dimana distribusi χ^2 dengan 1 derajat bebas dan ε adalah nilai *error* dari hasil *Ordinary Least Square*. Jika LM_{error} signifikan maka model yang sesuai adalah SEM.

2.11 Matriks Pembobot Spasial

Matriks *contiguity* adalah matriks yang menggambarkan hubungan antar wilayah pengamatan. Elemen W_{ij} akan memberikan nilai 1 pada wilayah yang berbatasan langsung dengan lokasi pengamatan dan sisanya diberikan nilai 0.

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

Untuk menentukan bagaimana hubungan spasial antar daerah pengamatan, dapat menggunakan berbagai metode dasar antara lain meliputi:

- a. *Queen Contiguity*

Hubungan kedekatannya berdasarkan pada langkah ratu pada papan catur. Daerah yang berhimpit kearah kanan, kiri, atas, bawah, dan diagonal didefinisikan sebagai daerah yang saling berdekatan.

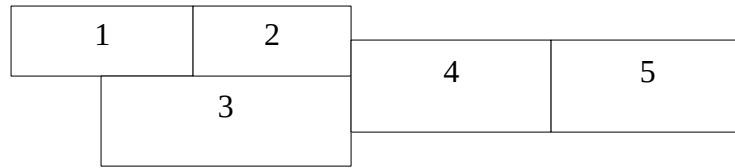
b. *Bishop Contiguity*

Hubungan spasial antar daerah pengamatan hanya dapat ditentukan dalam arah diagonal saja berdasarkan titik yang saling bersinggungan.

c. *Rook Contiguity*

Hubungan spasial antar daerah pengamatan dapat ditentukan berdasarkan sisi yang saling bersinggungan. Daerah pengamatannya dapat ditentukan kearah kanan, kiri, atas, dan bawah. Sedangkan arah diagonal tidak dapat ditentukan.

Penelitian ini menggunakan matriks *contiguity* dengan persinggungan sisi dan sudut (*Queen Contiguity*). Berikut ilustrasi contoh pembentukan matriks pembobot *Queen Contiguity*:



Gambar 2.2 Ilustrasi *Queen Contiguity*

Dari gambar 2.2 diatas, diperoleh matriks pembobot spasial berukuran 5×5 sebagai berikut:

$$W_{ij} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Baris dan kolom pada matriks menunjukkan wilayah yang ada pada peta. Elemen diagonal utama dari matriks W_{ij} adalah 0 pada wilayah yang tidak dianggap tetangga yang merupakan wilayah itu sendiri. Setelah mendapatkan wilayah tersebut selanjutnya matriks tersebut distandarisasi dengan cara membagi setiap elemen matriks dengan jumlah setiap baris pada matriks yang sudah distandarisasikan sama dengan satu. Rumus untuk standarisasi matriks *Queen Contiguity* sebagai berikut:

$$W_{ij(std)} = \frac{w_{ij}}{\sum_{i=j}^n w_{ij}} \quad (2.38)$$

Selanjutnya diperoleh matriks *Queen Contiguity* yang telah distandarisasi

$$W_{ij(std)} = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Matriks pembobot *Queen Contiguity* yang sudah distandarisasi inilah yang digunakan dalam metode-metode pengujian dan pemodelan pada analisis spasial berbasis area.

2.12 Pemilihan Model Terbaik

Pemilihan model terbaik dapat dilakukan dengan menggunakan *Akaike Information Criterion* (AIC) dan koefisien determinasi (R^2).

2.12.1 Akaike Information Criterion (AIC)

Pemilihan model terbaik berdasarkan nilai AIC digunakan untuk memperoleh model yang paling sesuai dengan variabel data yang digunakan. Model regresi terbaik adalah model regresi yang memiliki nilai AIC terkecil (Arifani, 2023). Rumus untuk menghitung AIC dinotasikan sebagai berikut:

$$AIC = -2Lm + 2m \quad (2.39)$$

Dimana:

Lm : Maksimum *log-likelihood*

m : Jumlah parameter model

2.12.2 Koefisien Determinasi (R^2)

Koefisien determinasi digunakan untuk memperoleh informasi mengenai kecocokan dari suatu model, dimana variasi yang terjadi pada variabel dependen dapat dijelaskan oleh variabel independen. Koefisien determinasi dilambangkan dengan R^2 dan umumnya dinyatakan dalam persentase. Besarnya nilai koefisien determinasi berkisar dari 0 hingga 1. Semakin besar nilai R^2 , maka model yang digunakan semakin baik (Kurniawan & Yuniarto, 2016).

$$R^2 = \frac{JKR}{JKT} = 1 - \frac{JKE}{JKT}$$

2.13 Penelitian Terdahulu

1. Hasil penelitian Risa Khairia Sari (2023)

Penelitian Risa Khairia Sari (2023), berjudul “*Pemodelan Faktor-Faktor Yang Mempengaruhi Stunting di Pekanbaru Dengan Pendekatan Spatial Autoregressive Model (SAR)*”. Penelitian ini merupakan penelitian yang menggunakan metode *Spatial Autoregressive (SAR)*. Penelitian ini bertujuan untuk memodelkan kejadian stunting di Kota Pekanbaru dan melihat faktor-faktor yang mempengaruhi kejadian stunting. Dari hasil analisis SAR menunjukkan terdapat tiga variabel yang signifikan terhadap kejadian stunting di Kota Pekanbaru yaitu inisiasi menyusui dini, ibu hamil mendapat Fe dan ibu hamil KEK. Dengan hasil koefisien determinasi model regresi klasik sebesar 71,07% dan koefisien determinasi model SAR sebesar 85,22%. Artinya model SAR lebih bagus dan variabel inisiasi menyusui dini, ibu hamil mendapat Fe dan ibu hamil KEK menjadi faktor penyebab kejadian stunting sebesar 85,22%. Sedangkan sisanya dipengaruhi oleh faktor lain diluar model.

2. Hasil penelitian Windy David Revildy, Siti Sarah Sobariah Lestarin Yollanda Nalita (2018)

Penelitian Windy David Revildy, Siti Sarah Sobariah Lestarin Yollanda Nalita (2018), berjudul “*Pemodelan Spatial Error Model (SEM) Angka Prevalensi Balita Pendek (Stunting) di Indonesia Tahun 2018*”. Metode yang digunakan adalah *Spatial Error Model (SEM)*”. Hasil pengujian menunjukkan prevalensi balita pendek di suatu wilayah dipengaruhi oleh RLS, IMD, ASI, TPM, dan sanitasi layak yang signifikan pada tingkat 10%. Model SEM menghasilkan nilai AIC sebesar 165,11 yang dinilai lebih baik dibanding model regresi metode OLS.

3. Hasil penelitian Atika Ismah Arifani (2023)

Penelitian Atika Ismah Arifani (2023), berjudul “*Perbandingan Spatial Autoregressive dan Spatial Error Model Dalam Pemodelan Indeks Pembangunan Manusia di Provinsi Sumatera Barat*”. Metode analisis yang digunakan dalam penelitian ini adalah metode Regresi Klasik dan Spatial Autoregressive (SAR). Penelitian ini bertujuan untuk melakukan pemodelan pada faktor yang berpengaruh terhadap IPM di Provinsi Sumatera Barat. Unit pengamatan pada penelitian ini adalah 19 kabupaten/kota di Provinsi Sumatera Barat. Hasil penelitian menunjukkan keempat faktor tersebut berpengaruh signifikan terhadap Indeks Pembangunan Manusia. Model Spatial Autoregressive memiliki AIC yaitu -23,5460 yang lebih kecil daripada AIC model Regresi Klasik sebesar 11.0465, artinya model *Spatial Autoregressive* adalah model terbaik untuk pemodelan indeks pembangunan manusia di Provinsi Sumatera Barat tahun 2020.

Tabel 2.3 Persamaan dan Perbedaan Penelitian Terdahulu

No.	Penelitian dan Tahun	Judul Penelitian	Persamaan	Perbedaan
1.	Risa Khairia Sari (2023)	Pemodelan Faktor-Faktor Yang Mempengaruhi Stunting di Pekanbaru Dengan Pendekatan Spatial Autoregressive Model (SAR)	Menggunakan data stunting Membahas mengenai Model SAR	Penelitian terdahulu mengambil kasus penelitian di Kota Pekanbaru sedangkan peneliti mengambil kasus di Provinsi Papua
2.	Windy David Revildy, Siti Sarah Sobariah Lestarin Yollanda	Pemodelan Spatial Error Model (SEM) Angka Prevalensi Balita Pendek (Stunting) di Indonesia Tahun 2018	Menggunakan data stunting Membahas mengenai Model SEM	Penelitian terdahulu mengambil kasus penelitian di Indonesia tahun 2018 sedangkan peneliti mengambil

No.	Penelitian dan Tahun	Judul Penelitian	Persamaan	Perbedaan
	Nalita (2018)			kasus di Provinsi Papua tahun 2022.
3.	Atika Ismah Arifani (2023)	Perbandingan Spatial Autoregressive dan Spatial <i>Error</i> Model Dalam Pemodelan Indeks Pembangunan Manusia di Provinsi Sumatera Barat	Membahas mengenai model SAR dan SEM	Peneliti terdahulu membahas kasus Indeks Pembangunan Manusia di Provinsi Sumatera Barat sedangkan peneliti membahas kasus stunting di Provinsi Papua.

2.14 Prevalensi Stunting

Stunting adalah suatu kondisi yang menggambarkan status gizi kurang yang memiliki sifat kronis pada masa pertumbuhan dan perkembangan anak. Dimulai sejak awal masa kehidupan yang dipresentasikan dengan *z-score* tinggi badan menurut umur kurang dari -2 standar deviasi berdasarkan standar pertumbuhan menurut WHO. Stunting merupakan kondisi yang disebabkan oleh kurang seimbangnya asupan nutrisi gizi pada masa pertumbuhan, bukan disebabkan oleh kelainan hormon pertumbuhan maupun akibat dari penyakit tertentu (Ni'mah & Nadhiroh, 2015).

Menurut Peraturan Menteri Kesehatan dalam Permenkes No. 2 Tahun 2020 tentang Standar Antropometri Penilaian Status Gizi Anak, pengertian pendek dan sangat pendek adalah status gizi yang didasarkan pada indeks antropometri. Indeks antropometri terdiri dari berat badan menurut umur (BB/U), tinggi badan menurut umur (TB/U), dan berat badan menurut tinggi badan (BB/TB). Indeks yang digunakan untuk mengetahui status balita stunting atau tidak adalah tinggi badan menurut umur (TB/U). Peraturan Menteri Kesehatan No. 2 Tahun 2020 mengenai Indeks Antropometri mengklasifikasikan status gizi stunting berdasarkan tinggi badan/panjang badan menurut umur sebagai berikut :

Tabel 2.4 Kategori dan Ambang Batas Status Gizi Anak

Indeks	Kategori Status Gizi	Ambang Batas (Z-score)
Panjang Badan atau Tinggi Badan menurut Umur (PB/U atau TB/U) anak usia 0 – 60 bulan	Sangat pendek	$< -3 SD$
	Pendek	$-3 SD \leq < -2 SD$
	Normal	$-2 SD \leq < +3 SD$
	Tinggi	$> +3 SD$

Umur (bulan)	Panjang Badan (cm)						
	-3 SD	-2 SD	-1 SD	Median	+1 SD	+2 SD	+3 SD
0	44.2	46.1	48.0	49.9	51.8	53.7	55.6
1	48.9	50.8	52.8	54.7	56.7	58.6	60.6
2	52.4	54.4	56.4	58.4	60.4	62.4	64.4
3	55.3	57.3	59.4	61.4	63.5	65.5	67.6
4	57.6	59.7	61.8	63.9	66.0	68.0	70.1
5	59.6	61.7	63.8	65.9	68.0	70.1	72.2

Gambar 2.3 Tabel Standar Antropometri

Standar Antropometri wajib digunakan sebagai acuan terkait penilaian status gizi anak. Berdasarkan gambar 2.3, jika pada usia 5 bulan seorang bayi memiliki panjang badan kurang dari 65,9 cm anak tersebut akan tergolong kategori status gizi normal menuju pendek. Kategori bayi pada usia 5 bulan dengan rentang panjang badan kurang dari 61,7 cm akan tergolong dalam kategori status gizi pendek dan untuk kategori status gizi sangat pendek akan masuk dalam rentang panjang badan kurang dari 59,6 cm.

Metode perhitungan prevalensi stunting menurut BAPPENAS (2020) adalah sebagai berikut :

$$PAB(5)P_{stunting} = \frac{JAB(5)P_{stunting}}{JAB(5)} \times 100\%$$

Dimana

PAB(5)Pstunting : Prevalensi anak balita yang menderita pendek (stunting)

JAB(5)Pstunting : Jumlah anak balita pendek (stunting) pada waktu tertentu

JAB(5) : Jumlah anak balita pada waktu yang sama

2.15 Faktor-Faktor Yang Mempengaruhi Stunting

2.15.1 Penduduk Miskin

Kemiskinan menjadi salah satu penyebab stunting yang dominan. Keterbatasan ekonomi membuat masyarakat tidak sanggup memenuhi kebutuhan gizi yang seimbang. Kerja keras untuk memperoleh penghasilan yang layak pun berujung pada pola asuh yang salah yang bisa berkontribusi pada stunting.

Pada penelitian Ramadhan (2018) dan Rahmawati (2020) banyaknya penduduk yang berstatus miskin akan mempengaruhi status gizi balita karena disebabkan oleh kurangnya masukan makanan yang bergizi sehingga dapat menyebabkan bertambahnya stunting di Indonesia.

Rumus perhitungan Persentase Penduduk Miskin menurut BPS :

$$P_a = \frac{1}{n} \sum_{i=1}^q \left[\frac{z - y_i}{z} \right]^a$$

Dimana :

a : 0

z : garis kemiskinan

y_i : Rata-rata pengeluaran per kapita sebulan penduduk yang berada dibawah garis kemiskinan ($i = 1, 2, 3, \dots, q$), $y_i < z$

q : Banyaknya penduduk yang berada di bawah garis kemiskinan.

n : Jumlah penduduk

2.15.2 Imunisasi Dasar lengkap

Imunisasi adalah suatu cara untuk meningkatkan kekebalan seseorang secara aktif terhadap antigen sehingga bila kelak terpajan pada antigen yang serupa tidak

terjadi penyakit. Imunisasi merupakan proses menginduksi imunitas secara buatan baik dengan vaksinasi (imunitas aktif) maupun dengan pemberian antibodi (imunisasi pasif). Dalam hal ini, imunisasi aktif menstimulasi sistem imun untuk membentuk antibodi dan respon imun seluler yang dapat melawan agen penginfeksi. Lain halnya dengan imunisasi pasif, imunisasi ini menyediakan proteksi sementara melalui pemberian antibodi yang diproduksi secara eksogen maupun transmisi transplasenta dari ibu ke janin (Ranuh dkk., 2008).

2.15.3 Pemberian ASI

Salah satu faktor penyebab stunting adalah tidak adekuatnya nutrisi pada masa pertumbuhan. Dalam masa pertumbuhan ASI berperan sangat penting dalam pertumbuhan anak, ketika nutrisi ini tidak dicukupi secara adekuat maka akan berpengaruh tinggi pada pertumbuhan dan perkembangan anak. Yang nantinya akan menghambat tinggi badan yang tidak sesuai dengan usianya (berperawakan pendek) (Prihutama dkk., 2018).

Memberikan ASI yang optimal kepada balita tidak hanya memenuhi kebutuhan gizi yang lengkap tetapi juga memberikan perlindungan terhadap penyakit karena ASI dapat meningkatkan Sistem kekebalan tubuh balita terhadap penyakit (Afifah, 2007).

2.15.4 Sanitasi Layak dan Sumber Air Bersih

Berdasarkan Riskesdas 2013, jika kondisi kesehatan lingkungan dikaitkan dengan status gizi, terutama prevalensi stunting pada balita maka terlihat bahwa masalah stunting sangat berhubungan dengan faktor kesehatan lingkungan yang tidak baik atau prevalensi stunting akan berkurang jika kondisi lingkungannya baik. Perbaikan akses sanitasi dan penyediaan air bersih akan menurunkan masalah stunting pada balita sebesar 20,58% (Kemenkes, 2013).

2.15.5 Faktor Pelayanan Kesehatan

Riskesdas 2013 menghasilkan indikator kesehatan dalam bentuk IPKM (Indeks Pembangunan Kesehatan Masyarakat). Pada indeks pelayanan kesehatan, yang merupakan komposit dari beberapa indikator, yaitu i) Persalinan oleh tenaga

kesehatan di fasilitas kesehatan, ii) Kecukupan jumlah dokter di setiap kecamatan, iii) Kecukupan jumlah posyandu, iv) kecukupan jumlah bidan v) kepemilikan jaminan pelayanan kesehatan dari penduduk. Dan hasilnya dapat disimpulkan bahwa perbaikan indeks pelayanan kesehatan dengan memperhatikan setiap masing-masing indikator dapat mengurangi masalah stunting pada balita sebesar 27.39%-36.29% (Kemenkes, 2013).

2.15.6 Indeks Ketahanan Pangan

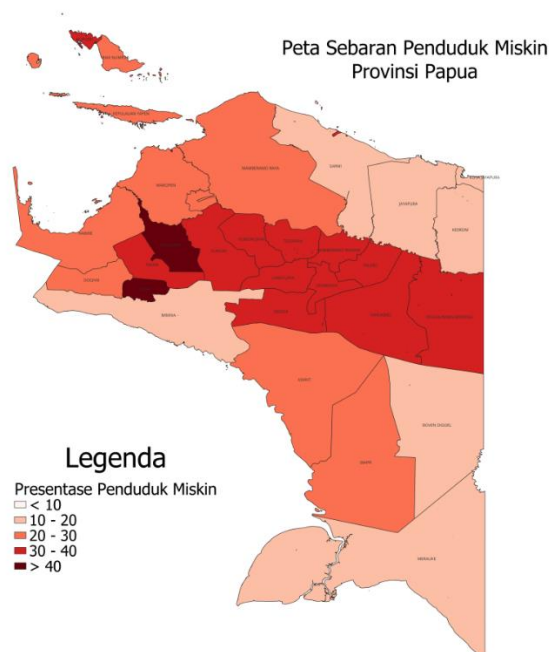
Ketersediaan pangan merupakan hal yang sangat penting karena termasuk dalam salah satu aspek dari ketahanan pangan. Ketahanan pangan bisa dijelaskan sebagai kemampuan seseorang atau keluarga untuk memenuhi kebutuhan pangan yang cukup baik dari segi kualitas maupun kuantitas, serta mampu mempertahankan akses tersebut secara berkelanjutan. Jika ketahanan pangan, terutama di tingkat rumah tangga, tidak mencukupi, maka dapat menyebabkan asupan pangan yang rendah dan berdampak buruk pada status gizi seseorang atau anggota keluarga.

Berdasarkan hasil penelitian Islamiah dkk. (2022) mendapatkan kesimpulan stunting pada balita terjadi karena faktor karakteristik keluarga, karakteristik balita, dan ketahanan pangan. Dan merekomendasikan adanya peningkatan ketahanan pangan dengan memperhatikan ketercukupan kebutuhan pangan keluarga serta memperhatikan pengeluaran rumah tangga dengan lebih mementingkan pemenuhan gizi balita dan meningkatkan ketahanan pangan rumah tangga dengan memberikan pelatihan dan keterampilan usaha.

2.16 Peta Tematik

Peta tematik adalah suatu bentuk peta yang menyajikan unsur-unsur tertentu dari permukaan bumi sesuai dari tema atau topik dari peta yang bersangkutan. Peta tematik umumnya digunakan sebagai data analisis dari beberapa unsur permukaan bumi di dalam pengambilan keputusan. Pada pembuatan peta tematik, peta topografi sebagai dasar sedangkan data tematik yang di sajikan adalah hasil survei langsung dan survei tidak langsung (Robinson dkk., 1995).

Pembagian kelas pada peta adalah teknik yang umumnya digunakan dalam pemetaan untuk menunjukkan varians atau pola sebaran data spasial dari suatu fenomena. Membagi data dalam tiga kelas saat membuat peta tematik adalah pilihan strategis untuk menyederhanakan visualisasi, membuat peta lebih mudah dipahami, dan efektif untuk menyampaikan informasi dasar atau perbedaan (Slocum dkk., 2022).



Gambar 2.4 Contoh Peta Tematik